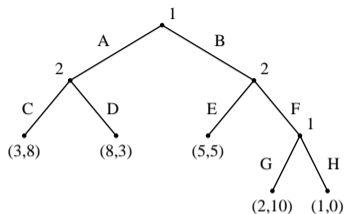




Subgame Perfection

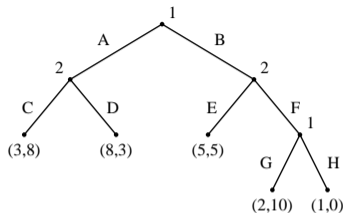
Game Theory Course:
Jackson, Leyton-Brown & Shoham

Subgame Perfection



- There's something **intuitively wrong** with the equilibrium $(B, H), (C, E)$
 - Why would player 1 ever choose to play H if he got to the second choice node?
 - After all, G dominates H for him

Subgame Perfection



- There's something **intuitively wrong** with the equilibrium $(B, H), (C, E)$
 - Why would player 1 ever choose to play H if he got to the second choice node?
 - After all, G dominates H for him
 - He does it to **threaten** player 2, to prevent him from choosing F , and so gets 5
 - However, this seems like a non-credible threat
 - If player 1 reached his second decision node, would he really follow through and play H ?

Formal Definition



Definition (subgame of G rooted at h)

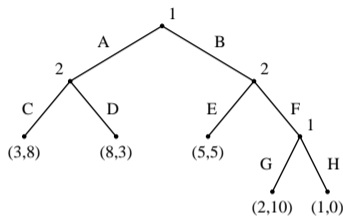
The **subgame of G rooted at h** is the restriction of G to the descendants of h .

Definition (subgames of G)

The **set of subgames of G** is defined by the subgames of G rooted at each of the nodes in G .

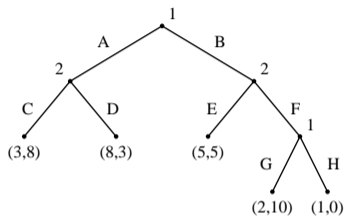
- s is a **subgame perfect equilibrium** of G iff for any subgame G' of G , the restriction of s to G' is a Nash equilibrium of G'
- Notes:
 - since G is its own subgame, every SPE is a NE.
 - this definition rules out “non-credible threats”

Which equilibria are subgame perfect?



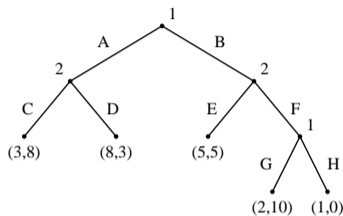
- Which equilibria from the example are subgame perfect?
 - $(A, G), (C, F)$:
 - $(B, H), (C, E)$:
 - $(A, H), (C, F)$:

Which equilibria are subgame perfect?



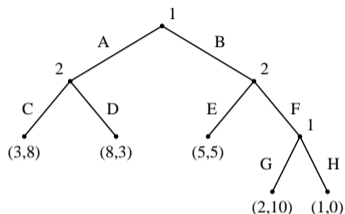
- Which equilibria from the example are subgame perfect?
 - $(A, G), (C, F)$: is subgame perfect
 - $(B, H), (C, E)$:
 - $(A, H), (C, F)$:

Which equilibria are subgame perfect?



- Which equilibria from the example are subgame perfect?
 - $(A, G), (C, F)$: is subgame perfect
 - $(B, H), (C, E)$: (B, H) is non-credible; not subgame perfect
 - $(A, H), (C, F)$:

Which equilibria are subgame perfect?



- Which equilibria from the example are subgame perfect?
 - $(A, G), (C, F)$: is subgame perfect
 - $(B, H), (C, E)$: (B, H) is non-credible; not subgame perfect
 - $(A, H), (C, F)$: (A, H) is non-credible, though H is “off-path”